Analysis of Mean Field Annealing in Subtractive Interference Cancellation

Thomas Fabricius *Student Member IEEE* and Ole Winther

tf@isp.imm.dtu.dk and owi@imm.dtu.dk

Informatics and Mathematical Modelling

Technical University of Denmark

2000 Lyngby, Denmark

Fax: +45 45 87 29 99

Phone: +45 45 25 38 94

Abstract

In this contribution we derive the cost function corresponding to the linear complexity Subtractive Interference Cancellation with tangent hyperbolic tentative decisions. We use the cost function to analyse the fix-points of solving the Subtractive Interference Cancellation equations. The analysis show that we can control the slope of the tangent hyperbolic functions so that the corresponding cost function is convex. We also show that increasing the slope can make the cost non-convex. Going from the convex regime into the non-convex regime, we prove that the bifurcation of the fix-points, for non-singular signal correlation matrices, consist of the fix-point of interest together with a saddle node bifurcation. This proves that tracking the solution from low slopes, with a convex cost, gradually increasing to higher slopes, with non-convex cost, can bring us to the best solution being very close to the optimal determined by enumeration. This tracking is the idea behind annealing. We show Monte Carlo studies with a substantial signal to noise ratio gain compared to not using annealing. We also show how annealing can be used to increase capacity at a given target bit error rate. In fact this capacity gain is the same obtained by *Improved Parallel Interference Cancellation* making us believe that the latter includes a mechanism to avoid local minima similar to annealing.

Index Terms

Multiple Access Technique, Subtractive Interference Cancellation, Local Minima, Bifurcation, Mean-Field Annealing, fix-point analysis

I. INTRODUCTION

During the last decades much effort has been used in inventing and developing multiuser communication systems. The ever growing demand for higher network capacity, to support more users, call for efficient communication systems which at the same time has low complexity. Multiple access schemes are traditional designed so that the individual users can be received independently, yielding low complexity and low cost receivers. In these schemes multiple access interference is kept at a low level, but so is the spectral efficiency. When increasing the spectral efficiency, the multiple access interference also increases. The optimal receiver is then based on a model taking the interference into account [1]. In the design there will exist a tradeoff between capacity measured in spectral efficiency and complexity measured as computations per information bit. This tradeoff has been the seed for many active research fields. Amongst these are multiuser detection (MUD), which today counts thousands of contributions. The optimal detector is for general systems exponential in the number of users, for a reference in Code Division Multiple Access (CDMA) see [2]. In order to realise such systems, suboptimal receivers are thus necessary.

When designing such a suboptimal receiver many factors influence the performance. When nuisance parameters, such as channel state including synchronisation, phase, frequency, signal to noise ratio, multiuser correlation etc., are not known a priori, the effect of estimating and tracking these can degrade the performance, especially in vehicular environments where fading phenomenons makes this a difficult task. Even when nuisance parameters are known exactly, the suboptimal detector will in some sense be an approximation to the optimal one, hence approximation errors degrade the performance. Many suboptimal detectors have a multistage or equivalent iterative nature, making convergence and local minima an issue in the performance degradation. Quantising the different sources to performance degradation is in general a difficult task. It is well accepted that using perfect knowledge about the nuisance parameters can be assumed in order to separate estimation and tracking errors from other errors, though there exists a complicated interaction. We will assume perfect nuisance parameter knowledge. For methods on estimation and tracking with semiblind or blind application consult [3], [4], [5], [6].

Suboptimal multiuser detectors are traditionally separated into two main categories [7]: Linear Detectors and Subtractive interference cancellation. Recently combinatorial optimisation methods have received attention [8], [9]. We also mention the earlier work [10] which maps CDMA multiuser detection to a Hopfield Neural Network, which is a general method to approximatively solve hard combinatorial problems [11]. In Ref. [9] binary constraints of the combinatorial optimisation are erlaxed: no constraints lead to the unconstrained Maximum Likelihood (ML) detection equivalent to the linear decorrelating/zero forcing (ZF) detector, and for other constraints they obtain subtractive interference cancellators with corresponding tentative decision functions. This work follows along the same lines. However, the constraints are derived by a certain Kullback-Leibler ($\mathcal{KL}$) divergence, which also provides us with a cost function.

A. Subtractive Interference Cancellation

Subtractive interference cancellation covers a whole family of multiuser detectors. The canonical form for the fix-point conditions in these detectors are given by

$$m_k^* = f\left(z_k - I_k(\mathbf{m}^*)\right) \qquad k = 1, \ldots, K \tag{1}$$

where z_k is the received data after the k'th matched filter, m_k^* is the tentative decision of the k'th symbol, $\mathbf{m}^*$ the tentative decisions of all bits arranged in a vector and I_k is the reconstructed interference on the k'th symbol's decision statistic, $f(x)$ is some tentative decision function which for binary antipodal symbols has asymptotes ± 1 for $x \to \pm\infty$. The equations (1) have been applied for various kinds of interference: Inter Symbol Interference (ISI) environment [12] and for CDMA Multiuser Interference [13], [14]. The equations can be solved sequentially or in parallel, leading to Serial or Parallel Interference Cancellation (PIC or SIC). Sequential updates are in most cases more numerical stable, but introduces an extra processing delay compared to parallel update which unfortunately can have an oscillating behaviour [15]. We also mention *Improved Parallel Interference Cancellation* [16] and the work using *Probabilistic Data Association* [17]. We postulate that these methods include a mechanism for avoiding local minima that are more or less equivalent to *annealing*, which we introduce in this contribution, we discuss this in section IX. The first mention of the possibility of using annealing as a heuristic in CDMA was made by Tanaka [18].

Various tentative decision functions are suggested in the literature including sign, hyperbolic tangent, clipped soft decision, etc. [16]. The first corresponding to maximal constrained ML, the last to a box constraint [9], [19]. Hyperbolic tangent is well-known to be optimal in the absence of interference. We will concentrate on subtractive interference cancellation with hyperbolic tangent, since this is the function that follows from the $\mathcal{KL}$-divergence framework. We show in section IV how this approximative detector relates to the individual optimal detector when multi-access interference (MAI) is present, and which constraints it implies. Hyperbolic tangent subtractive interference cancellation, as we consider it, is previously derived in the physics literature on binary spin systems under the name Naive Mean Field (NMF) [20]. More advanced mean field approaches exist, which has a potentially lower approximation error [21], [22], [23]. Mean field methods has earlier been applied to Independent Component Analysis (ICA) [24], [25]. The equivalence between MUD and ICA has been exploited in [26].

Besides approximation error, local minima or bad convergence of solving (1) also degrade the performance. In section VI we therefore make a thorough analysis of the fix-points as function of the hyperbolic tangents slope. The analysis provides us very important information about the fix-points, when they are convex and how they bifurcate. The bifurcation analysis helps us to undertand the behaviour of local minima. Having got this understanding, makes us able to solve the equations (1) while avoiding local minima. This way of solving the equations is called mean field annealing, which we introduce in section VII. The analysis of the bifurcations validates the use of mean field annealing. Annealing is often used as a heuristic to avoid local minima. Annealing saw its first use in optimisation with the introduction of simulated annealing applied to the design of computer chips [27]. Mean field annealing has been used in various optimisation problems, for instance travelling salesman problem, graph partitioning, image restoration, etc. [11], [28]. Our analysis show that mean field annealing will occasionally fail no matter how fine an annealing scheme employed. But we show that this is the case where even the optimal detector can fail, leading to the fundamental bit error rate (BER) floor [19]. Large system analysis of subtractive interference cancellation's bit error rate also suggest the use of annealing, since competing unequivocal and indecisive minima, exists in this limit [18]. Thus in the large system limit the existence of two phases only is predicted for loads higher than unity, this is not the case for finite size systems.

We will assume a very simple model where the interesting effects can be studied, namely a CDMA system with K synchronous users employing Binary Phase Shift Keying (BPSK) with random binary spreading sequences. For an overview of MUD in CDMA consult [7]. We derive the most important result for any given power profile of the users, other results are given for equal power users which form a worst case scenario in terms of local minima. Most of the derived result reduces to rely on the matrix of users cross correlation, which could stem from any multiuser system and/or inter symbol interference, so we conjecture that the results hold for typical multiuser systems synchronous or asynchronous.

II. K USERS CDMA IN AWGN

We will assume a CDMA transmission employing binary phase shift keying (BPSK) symbols and using K random binary spreading codes with unit energy and equal spreading factor (SF) but different powers and all chip and symbol synchronised over an additive white Gaussian noise (AWGN) channel. Now the received base band CDMA signal can be modelled as

$$y(n) = \sum_{k=1}^{K} A_k b_k s_k(n) + \sigma\epsilon(n) \ , \tag{2}$$

where n is sample index $n \in [0, \ldots, N_c - 1]$, N_c being the number of chips, i.e. equalling the spreading factor, and $s_k(n)$ is the k'th users spreading code with unit energy, $b_k \in [-1, 1]$ is user k's transmitted bit, A_k the k'th users amplitude, $\epsilon(n)$ is a Gaussian white noise sample with zero mean and variance 1, and σ is the noise standard deviation. Making this normalisation makes us able to define the k'th users signal to noise ratio as $SNR_k = \frac{A_k^2}{2\sigma^2}$.

We correlate the signal $y(n)$ with the spreading codes $s'_k(n)$ each code denoted by $k' \in [1; K]$, to obtain the conventional detector outputs $z_{k'}$ which are sufficient statistics for the symbol estimation

$$z_{k'} = \sum_{n=0}^{N_c-1} y(n) s_{k'}(n) = \sum_{k=1}^{K} A_k b_k \sum_{n=0}^{N_c-1} s_{k'}(n) s_k(n) + \sigma \sum_{n=0}^{N_c-1} s_{k'}(n)\epsilon(n) = \sum_{k=1}^{K} A_k b_k r_{kk'} + e_{k'} \ , \tag{3}$$

where $r_{kk'} = \sum_{n=0}^{N_c-1} s_{k'}(n) s_k(n)$ is the correlation between code s_k and $s_{k'}$, and $e_{k'}$ is now a Gaussian random variable with zero mean and covariance $\sigma^2 r_{kk'}$. Using the above we have the joint distribution of the $z_{k'}$

$$\begin{aligned} &p\left(\mathbf{z} \,|\mathbf{b}, \mathbf{A}, \mathbf{R}, \sigma^2\right) = \\ &|2\pi\sigma^2\mathbf{R}|^{-\frac{1}{2}} \exp - \left(\tfrac{1}{2\sigma^2}(\mathbf{z} - \mathbf{RAb})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{z} - \mathbf{RAb})\right) \quad , \end{aligned} \tag{4}$$

where we have arranged the vectors to have elements $(\mathbf{z})_{k'} = z_{k'}$ and $(\mathbf{b})_{k'} = b_{k'}$, and the matrices $(\mathbf{A})_{kk'} = \delta_{kk'} A_k$, and $(\mathbf{R})_{kk'} = r_{kk'}$. Since we have assumed perfect channel state knowledge, and we also assume full code knowledge the matrices $\mathbf{A}$ and $\mathbf{R}$ and the noise variance σ^2 are not considered stochastic. We then write the likelihood as $p\left(\mathbf{z} \,|\mathbf{b}\right)$ instead of $p\left(\mathbf{z} \,|\mathbf{b}, \mathbf{A}, \mathbf{R}, \sigma^2\right)$.

III. Review of Optimal Detectors

Given the received data and the channel, one may either try to minimise the expected bit error rate (BER) or the probability of error. The expected BER is defined as

$$\overline{\mathrm{BER}} = \frac{1}{2}\langle 1 - \frac{1}{K}\mathbf{b}^{\mathrm{T}}\widehat{\mathbf{b}}\rangle_{p(\mathbf{z},\mathbf{b})} \tag{5}$$

where the average $\langle\cdot\rangle_{p(\mathbf{z},\mathbf{b})}$, as indicated, is taken with respect to all bit realisations $\mathbf{b}$ and all noise realisations imbedded in $\mathbf{z}$. The joint distribution $p(\mathbf{z}, \mathbf{b})$ is found by the likelihood (4) multiplied by the a priori distribution $p(\mathbf{z}, \mathbf{b}) = p\left(\mathbf{z} \,|\mathbf{b}\right) p(\mathbf{b})$.

It can easily be shown, under some regularity conditions, that the expected BER subject to the constraint that $\left|(\widehat{\mathbf{b}})_{k'}\right| = 1$ is minimised by

$$\widehat{\mathbf{b}} = \mathrm{sgn}\langle \mathbf{b} \rangle_{p(\mathbf{b}|\mathbf{z})} \tag{6}$$

for all given received data $\mathbf{z}$, sgn working element-wise. This estimator is referred to as the individual optimal detector [29].

The probability of error is defined as the probability that at least one of the K bits are detected wrongly. We can define $p(e = 1\,|\mathbf{b}) = 1 - \delta(K - \mathbf{b}^{\mathrm{T}}\widehat{\mathbf{b}})$ and so the probability of error becomes

$$p(e = 1) = 1 - \langle \delta(K - \mathbf{b}^{\mathrm{T}}\widehat{\mathbf{b}}) \rangle_{p(\mathbf{z},\mathbf{b})} \tag{7}$$

which is minimised for any received data $\mathbf{z}$ [17] by

$$\widehat{\mathbf{b}} = \underset{\mathbf{b}\in[-1;1]^K}{\mathrm{argmax}}\, p(\mathbf{b}\,|\mathbf{z}) \tag{8}$$

which is the well-known Maximum A Posterior (MAP) solution, which reduces to maximum likelihood for uniform prior distribution $p(\mathbf{b})$, also denoted the joint optimal detector [29].

The two above optimal detectors deliver hard symbols as output. This is well-known not to be optimal if the symbols are sufficiently interleaved and the detection is followed by redundancy decoding. Then it is optimal to output the marginal symbol posterior. In case the symbols are binary, one parameter per binary symbol is sufficient to describe the corresponding marginal posterior. The parameter could be the marginal posterior mean $\langle \mathbf{b} \rangle_{p(\mathbf{b}|\mathbf{z})}$ or the log posterior ratio $\log \frac{p(1|\mathbf{z})}{p(-1|\mathbf{z})}$ which is identical to the log likelihood ratio under equal probable a priori symbols. We call this the soft individual optimal detector.

For the model under consideration in this contribution, all the above detectors, for general spreading codes, have exponential complexity in the number of users K. A big effort therefore is to construct approximative polynomial time complexity detectors.

IV. KL-DIVERGENCE BETWEEN OPTIMAL AND APPROXIMATIVE DETECTORS

All the optimal detectors rely heavily on the posterior over the symbols, either as the maximising argument or the mean. For that reason will we now consider a distribution that approximates the posterior distribution, but in which the maximising argument and the posterior mean is obtained in polynomial time. As a measure of the closeness of the approximating distribution, which we denote $q(\mathbf{b})$, to the posterior distribution $p(\mathbf{b}\,|\mathbf{z}\,)$ we use the Kullback-Leibler divergence

$$\mathcal{KL}(q(\mathbf{b}), p(\mathbf{b}\,|\mathbf{z}\,)) = \sum_{\mathbf{b}\in[-1,1]^K} q(\mathbf{b}) \log \frac{q(\mathbf{b})}{p(\mathbf{b}\,|\mathbf{z}\,)} = \langle \log \frac{q(\mathbf{b})}{p(\mathbf{b}\,|\mathbf{z}\,)} \rangle_{q(\mathbf{b})}. \tag{9}$$

We now take the definition of the posterior distribution using the likelihood and equal a priori symbol probability

$$p(\mathbf{b}\,|\mathbf{z}\,) = \frac{p(\mathbf{z}\,|\mathbf{b}\,)}{Z} \tag{10}$$

where $Z = \sum_{\mathbf{b}\in[-1,1]^K} p(\mathbf{z}\,|\mathbf{b}\,)$ is the partition function or normalising constant; and use this in the $\mathcal{KL}$-divergence

$$\begin{aligned}\mathcal{KL}(q(\mathbf{b}), p(\mathbf{b}\,|\mathbf{z}\,)) =& \frac{1}{2\sigma^2} \sum_{k,k'=1}^{K} A_k A_{k'} \langle b_k b_{k'} \rangle_{q(\mathbf{b})} (r_{kk'} - \delta_{kk'}) \\ & - \frac{1}{\sigma^2} \sum_{k=1}^{K} A_k z_k \langle b_k \rangle_{q(\mathbf{b})} + \langle \log q(\mathbf{b}) \rangle_{q(\mathbf{b})} + const.\end{aligned} \tag{11}$$

where the constant $const. = \frac{1}{2\sigma^2}(\mathbf{z}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{z} + \sum_{k=1}^{K} A_k^2) + \log Z + \frac{1}{2}\log|2\pi\sigma^2 R|$ is independent of the distribution $q(\mathbf{b})$. We now introduce a new function that contains the parts that depends on $q(\mathbf{b})$

$$\mathcal{F}(T, q(\mathbf{b})) = \frac{1}{2} \sum_{k,k'=1}^{K} A_k A_{k'} \langle b_k b_{k'} \rangle_{q(\mathbf{b})} (r_{kk'} - \delta_{kk'}) - \sum_{k=1}^{K} A_k z_k \langle b_k \rangle_{q(\mathbf{b})} + T \langle \log q(\mathbf{b}) \rangle_{q(\mathbf{b})}. \tag{12}$$

where we have substituted σ^2 with T so that we can study $\mathcal{F}$ at various T and not only on the generic $T = \sigma^2$ which is the noise level. We then have $\mathcal{KL}(q(\mathbf{b}), p(\mathbf{b}\,|\mathbf{z}\,)) = \frac{1}{T}\mathcal{F}(T, q(\mathbf{b}))\big|_{T=\sigma^2} + const.$ The choice of notation is influenced by statistical physics where the function $\mathcal{F}$ is denoted the *free energy* (up to a constant) and T corresponds to the temperature of the system.

We can consider $\mathcal{F}$ as a cost function for $q(\mathbf{b})$ at a given T. The optimum distribution $q(\mathbf{b})$ that minimises the $\mathcal{KL}$-divergence, and hence also the free energy, is obviously $q(\mathbf{b}) = p(\mathbf{b}\,|\mathbf{z}\,)$, but this of course does not have the desired properties, namely ease of obtaining the posterior maximising argument or posterior mean.

The idea is now clear, we can come up with *trial* distributions $q(\mathbf{b})$, eventually parameterised, and choose the one that minimises $\mathcal{F}$, and thereby the difference between the distributions. We now do the simplest possible assumption about the *dependence* relation of the distribution $q(\mathbf{b})$, namely independence[1]

$$q(\mathbf{b}) = \prod_{k=1}^{K} q_i\left(b_k\right) . \tag{13}$$

Only the family of Bernoulli distributions are suitable for the individual $q_k\left(b_k\right)$, since the b_k's are binary

$$q_k(b_k) = p_i^{\frac{1+b_k}{2}} (1 - p_k)^{\frac{1-b_k}{2}} \ , \tag{14}$$

where we have parameterised $q_k\left(b_k\right)$ by p_k the probability of $b_k = 1$. We will thus for mathematical convenience choose a parameterisations by the mean $m_k \equiv \langle b_k \rangle_{q_k} = 2p_k - 1$

The free energy then becomes

$$\mathcal{F}(T, q(\mathbf{b})) = \frac{1}{2} \sum_{k,k'=1}^{K} A_k A_{k'} m_k m_{k'} (r_{kk'} - \delta_{kk'}) - \sum_{k=1}^{K} A_k z_k m_k - T \sum_{k=1}^{K} \mathcal{H}\left(q_k\right) \ , \tag{15}$$

where we introduce the entropy of one such distribution $q_i\left(b_i\right)$

$$\mathcal{H}\left(q_i\right) \equiv -\langle \log q_k \rangle_{q_k} = -\frac{1+m_k}{2} \log \frac{1+m_k}{2} - \frac{1-m_k}{2} \log \frac{1-m_k}{2} \ . \tag{16}$$

When minimising equation (15), we see that the different terms compete. The entropy term should be maximised, which happen in $m_k = 0$, in order to minimise (15). We also see that the entropy permits values of $|m_k| > 1$. The remaining terms makes up $-\log$ likelihood which as usual should be minimised.

V. Optimising the Free Energy

Unfortunately it is not possible to minimise the free energy analytically due to the non-linear entropy term of the free energy. Numerical minimisation can then be applied. The method used will form the *dynamics* of the minimisation, i.e. the change of variables over time (iteration). Considering the minimisation as a *dynamical system*, we strive for the stable fix-point of the dynamics.

We will now introduce a fix-point method obtained from the stationary condition

$$\frac{\partial \mathcal{F}\left(T, q(\mathbf{x})\right)}{\partial \mathbf{m}} = 0 \ . \tag{17}$$

[1] It is of course naive to believe in independence when we now this is not true, for that reason the assumption is often denoted the *naive* assumption or approximation.

A straight forward calculation gives us the non-linear fix-point equations

$$m_k = \tanh\left(\frac{A_k}{T}\left(z_k - \sum_{k'=0}^{K-1} A_{k'}(r_{kk'} - \delta_{kk'})m_{k'}\right)\right) \tag{18}$$

for $k = 1, \ldots, K$. We see that this equation is the well-known soft decision feedback subtractive interference cancellation with a version of clipped soft decision function namely $\tanh(\cdot)$, where $I(\mathbf{m}) \equiv \sum_{k'=1}^{K} A_{k'}(r_{kk'} - \delta_{kk'})m_{k'}$ is the estimated interference correlated with the k'th users spreading code. Equalling T to the signal to noise ratio, the understanding of equation (18) is straight forward because in the absence of multi-access interference (MAI) this is the optimal decision in an AWGN channel. This is clear, because absence of MAI is identical to the naive assumption being exact, namely the true posterior factorises. We also have

$$\begin{aligned} m_k &= \lim_{T\to 0} \tanh\left(\frac{A_k}{T}\left(z_k - \sum_{k'=1}^{K} A_{k'}(r_{kk'} - \delta_{kk'})m_{k'}\right)\right) && (19) \\ &= \operatorname{sgn}\left(A_k\left(z_k - \sum_{k'=1}^{K} A_{k'}(r_{kk'} - \delta_{kk'})m_{k'}\right)\right) . && (20) \end{aligned}$$

This is the hard decision feedback interference cancellation, known to be a local search instance of Maximum Likelihood sequence estimation [9].

Beside our two contributions [6], [30], the connection between the equations (18) and the minimisation of a certain cost function (the free energy), is to our knowledge never established in the communication literature[2]. Though in physics and machine learning this approach is well-known as the *variational approach* and the restriction to a factorised trial distribution is referred the *naive assumption*, the corresponding free energy is referred the *naive free energy*, and the equations (18) is denoted the naive mean-field equations [23]. The benefit of having the corresponding cost function to the update rules are numerous: We can get an understanding of the quality of approximation, i.e. when it is exact. We can use it to measure the likelihood of the fix-points, which is useful for defining stopping criteria and for line search updates. Also

[2]But a similar connection has been established for the approximative updates making up the turbo principle to a certain free energy, the *Bethe free energy* [31]

note that sequential update of equation (18) corresponding to coordinate descent is guaranteed to decrease the free energy since the update of m_k does not depend on itself. In this contribution we investigate various aspects of the free energy, i.e. convexity, uniqueness of fix-point solutions, and fix-point bifurcations as the slope of the tentative decision functions change.

VI. Study of Fix-points in the naive free energy

Subtractive interference cancellation with decision feedback is known to suffer from error propagation, which is best described as an *avalanche* effect from one erroneous detected symbol on other symbols due to their correlation. This error propagation corresponds to different fix-points of the free energy. It is also known that the softer the mapping function is, here controlled by T, the fewer stable fix-points exists, indeed the infinitely soft mapping function is linear, yielding a convex optimisation problem. We will firstly analyse the convexity of the free energy as function of T the slope of the mapping function, but also at which T we can expect the fix-points to bifurcate. We will later analyse the type of fix-points as function of T for given received signal $\mathbf{z}$ and correlation matrix $\mathbf{R}$.

A. Convex fix-point analysis

We will first address how high a T that is needed for the free energy to posses only one fix-point. The critical T where this happens will we denote T_{c_1}. Below T_{c_1} we can expect a bifurcation of the fix-point to multiple fix-points.

The sufficient and necessary requirement for the free energy to posses one fix-point is that all eigenvalues of the Hessian matrix to be positive in the whole phase space $\mathbf{m}$ inside the hyper cube $[-1;1]^K$. The Hessian being $(\mathbf{H}(\mathbf{m}))_{ij} = \frac{\partial^2}{\partial m_i \partial m_j}\mathcal{F}(T, \mathbf{h}, q(\mathbf{x})) \implies \mathbf{H}(\mathbf{m}) = \mathbf{R} - \mathbf{I} + T\mathbf{W}(\mathbf{m})$ and $(\mathbf{W}(\mathbf{m}))_{ij} = \delta_{ij}\frac{1}{1-(\mathbf{m})_i^2}$ being the second derivative of the negative entropy. Using this requirement we obtain

Theorem 1 (One Solution). *Given the form (15) of the free energy, the smallest eigenvalue $\rho_{\min}$ of the code correlation matrix $\mathbf{R}$, the maximum user power $A_{\max}^2 \equiv \max_i A_i^2$ then the free energy is strictly convex for*

$$T > T_{c_1} = (1 - \rho_{\min})\, A_{\max}^2 \tag{21}$$

Proof: Let $\lambda_1 \leq \ldots \leq \lambda_K$ be an ascending ordering of the Hessian eigenvalues then the requirement for positiveness is equivalent to requiring $\lambda_1 > 0$. We bound the smallest eigenvalue

$$\begin{aligned}\lambda_1 &= \min \operatorname{eig} \mathbf{A}(\mathbf{R}-\mathbf{I})\mathbf{A} + T\mathbf{W}(\mathbf{m}) \geq \min \operatorname{eig} \mathbf{A}(\mathbf{R}-\mathbf{I})\mathbf{A} + T\mathbf{W}(\mathbf{0}) \\ &= \min \operatorname{eig} \mathbf{A}(\mathbf{R}-\mathbf{I})\mathbf{A} + T\mathbf{I} \geq (\rho_{\min}-1) \max_i A_i^2 + T \,,\end{aligned} \tag{22}$$

where we used that the the smallest eigenvalue $\rho_{\min} - 1$ of $\mathbf{R} - \mathbf{I}$ is always less or equal to zero, since $\sum_{k=1}^K \rho_k = \operatorname{tr} \mathbf{R} = K$. This implies the smallest eigenvalue of $\mathbf{A}(\mathbf{R}-\mathbf{I})\mathbf{A}$ to be greater than or equal to $(\rho_{\min} - 1)\, A_{\max}^2$. Our requirement for at most one solution is $\lambda_l > 0$ so if

$$T > (1-\rho_{\min})\, A_{\max}^2 \implies \lambda_l > (\rho_{\min}-1)\, A_{\max}^2 + T > 0 \,. \blacksquare \tag{23}$$

If $\rho_{\min} = 1$, we only have one minimum no matter the SNR's. But if $\rho_{\min} = 1$ then all the eigenvalues of the code correlation matrix equal one since $\sum_{k=1}^K \rho_k = \operatorname{tr} \mathbf{R} = K$. This notion corresponds to the code correlation matrix being diagonal since the *Frobenius norm* defined as $||\mathbf{R}||_F = \sqrt{\sum_{k,k'=1}^K \mathbf{R}}$ fulfills $\rho_{\max} \leq ||\mathbf{R}||_F \leq \sqrt{K}\rho_{\max}$ and hence the off-diagonal elements must equal zero. This implies that we have no interference and then the factorised assumption is exact. So not only do we surely have one and only one minimum, the approximation obtained by the $\mathcal{KL}$-divergence also yields the exact result! The theorem 1 has a conservative version

Corollary 1 (Conservative One Solution). *A conservative estimate of when the free energy posses one and only one minimum is obtained by*

$$T > A_{\max}^2 \,. \tag{24}$$

Proof: We have directly from the previous proof

$$T > (1-\rho_{\min})\, A_{\max}^2 \geq A_{\max}^2 \,, \tag{25}$$

where we use that the smallest eigenvalue over all code realisations $\rho_{\min} \geq 0$, since the code correlation matrix $\mathbf{R}$ is positive semi-definite. $\blacksquare$

The corollary offers the option to select T independent of the actual eigenvalues of the code correlation matrix; by choosing $T = A_{\max}^2$, which conservatively ensures the existence of one and only one fix-point, and we won't see error propagation.

We now have the opportunities to select T in order to ensure the existence of one and only one minimum. But if we choose so, then in general we have $T > \sigma^2$ corresponding to overestimating the noise and hence we can expect the power of the residual multi-access interference (MAI) (actual MAI minus estimated MAI) to be on the order $\mathcal{O}(T-\sigma^2)$. This means that we maybe cancelled less of the MAI than possible. We see this as an excess bit error rate in the asymptote with infinitely good SNR.

B. Bifurcation analysis

Now having analysed at which value of T where a bifurcation from one fix-point into more fix-points is expected, we want to understand the srtucture of the fix-points and how they relate when we change T. In order to obtain such an understanding we make a classical characterisation of this bifurcation, and later bifurcations. We will study the occurrence of two or more equivalent minima, where equivalent means exact same numerical value of the free energy. When equivalent minima are observed in the free energy, there is a direct correspondence to the fundamental BER-floor of the joint optimal detector, described in [19]. Equivalent minima occurs when one or more users signal exactly cancels. The joint optimal detector is implemented by enumeration, where the likelihood of all the 2^K possible combination are calculated. Then equivalent minima means that two or more of the enumerated values are identical and the detector has to choose at random between these. This implies that if the probability for such identical minima to occur is finite, then we will experience a BER-floor [19]. The same is observed for the optimal individual detector, where one or more of the symbol posterior mean values equal zero which reflect the ambiguity in the Likelihood due to the cancellation.

In the bifurcation analysis will we assume perfect power control, i.e. without loss of generality all $A_k^2 = 1$, since only under these conditions can two users exactly cancel each other. They furthermore need to have identical time aligned spreading codes.[3]. This implies that equal power profile and the synchronous model is worst case in terms of local minima!

We will assume that the code correlation matrix's eigenvalues are non-degenerate (multiplicity one) which is a reasonable assumption for moderate to high spreading factors N, since the eigenvalue spectrum tends towards a continuous spectrum. Non-degenerate code correlation eigenvalues generally implies that the Hessian's eigenvalues are also non-degenerate. Since the *bifurcation of a fix-point is determined in the zero eigen-directions of the Hessian* [32], the

[3]For non-perfect power control three or more users can cancel each other if their respective codes permits it, but this has a lower probability than the mentioned case.

previous assumption is equivalent to saying that we can consider the bifurcation in one direction which we denote $\mathbf{v}_{k'}$ with corresponding code correlation eigenvalue $\rho_{k'}$.

The main result of the bicurcation analysis, given in appendix I, is the so-called *bifurcation set* or *spinodal lines*

$$\left(\frac{3\mathbf{v}_{k'}^{\mathrm{T}}\mathbf{z}}{2T\sum_{i=1}^{K}(\mathbf{v}_{k'})_i^4}\right)^2 + \left(\frac{T - T_{c_{k'}}}{T\sum_{i=1}^{K}(\mathbf{v}_{k'})_i^4}\right)^3 = 0 \tag{26}$$

where $T_{c_{k'}} = 1 - \rho_{k'}$. The bifurcation set gives the conditions for going from a single fix-point solution (lhs > 0) to three non-degenerate solutions (lhs < 0) of which two are stable corresponding to a minimum of free energy and one is unstable. In figure 1, we plot the bifurcation set which gives rise to the *cusp* figure. Inside one cusp we have three fix-points. Outside the cusp there is only one fix-point. We have only drawn cusps corresponding to $0 \leq \frac{T}{3}\sum_{i=1}^{K}(\mathbf{v}_{k'})_i^4 \leq \frac{1}{3}$, since from corollary 1 we know that for $\sum_{i=1}^{K}(\mathbf{v}_{k'}^4)_i \leq 1$ and $T > 1$, only one solution exists. We also note that $\rho_{k'} - 1 + T \geq -1$ since $T \geq 0$ and $\rho_{k'} \geq 0$.

We will now look at the consequences of the bifurcation set. The condition for triple degeneracy is $\rho_{k'} - 1 + T = 0$ and $\mathbf{v}_{k'}^{\mathrm{T}}\mathbf{z} = 0$. The first fulfilled exactly at $T = T_{c_{k'}}$ below $T_{c_{k'}}$ we have three solutions $x_{k'} = 0$ which is a maximum and $x_{k'} = \pm\sqrt{\frac{3\left(T_{c_{k'}} - T\right)}{T\sum_{k=1}^{K}(\mathbf{v}_{k'})_k^4}}$ that are minima. This type of bifurcation is the classical pitchfork bifurcation, see top plots of figure 1, where one valid fix-point becomes three fix-points. The fix-point in the middle is unstable since the curvature is negative. The two other fix-points are minima since the curvature is positive. But they are also equivalent minima meaning that they have the same free energy, see figure 1 top right. This mean that any search method, even if we try all possible values, will at random choose either the wrong or the right one, the first corresponding to error propagation. Since this bifurcation is so critical that even the optimal detector can make errors, we want to examine the condition for $\mathbf{v}_{k'}^{\mathrm{T}}\mathbf{z} = 0$ in terms of the code correlation matrix and hence the existence of equivalent minima. We derive the condition for equivalent minima in one direction $\mathbf{v}_{k'}$

Theorem 2 (Equivalent minima in one direction). *Let $\mathbf{v}_{k'}$ be an eigenvector to the code correlation matrix $\mathbf{R}$, let $\rho_{k'}$ be the corresponding eigenvalue, and let $T \leq T_{c_{k'}}$ then a sufficient condition for two minima along $\mathbf{v}_{k'}$ to exist is $\rho_{k'} = 0$.*

Proof: We start by writing out the received and matched filtered received vector $\mathbf{z}$ in terms of the transmitted symbols $\mathbf{b}$, the code correlation matrix $\mathbf{R}$, and the coloured noise vector $\mathbf{e}$

with covariance $\sigma^2\mathbf{R}$

$$\mathbf{z} = \mathbf{R}\mathbf{b} + \mathbf{e} \tag{27}$$

then the projection onto $\mathbf{v}_{k'}$ is

$$\mathbf{v}_{k'}^{\mathrm{T}}\mathbf{z} = \mathbf{v}_{k'}^{\mathrm{T}}\mathbf{R}\mathbf{b} + \mathbf{v}_{k'}^{\mathrm{T}}\mathbf{e} = \rho_{k'}\mathbf{v}_{k'}^{\mathrm{T}}\mathbf{b} + \sqrt{\rho_{k'}}\mathbf{v}_{k'}^{\mathrm{T}}\sigma\mathbf{e}_w \ , \tag{28}$$

where we used $\mathbf{v}_{k'}^{\mathrm{T}}\mathbf{R} = \rho_{k'}\mathbf{v}_{k'}^{\mathrm{T}}$ and that $\mathbf{e} = \sum_{k=1}^{K}\mathbf{v}_k\sqrt{\rho_k}\mathbf{v}_k^{\mathrm{T}}\sigma\mathbf{e}_w$ where $\mathbf{e}_w$ is a K dimensional white standard Gaussian vector. Now if $\rho_{k'} = 0$ we have $\mathbf{v}_{k'}^{\mathrm{T}}\mathbf{z} = 0$ which was the condition for two equivalent minima when $T \leq T_{c_{k'}}$. Hence $\rho_{k'} = 0$ is sufficient for this to hold. ■

We see that a sufficient condition if $T \leq T_{c_{k'}}$ for two equivalent minima to occur is when the code correlation matrix is singular. This for instance happens if two or more users are assigned the same spreading code. Under random code assignment, the probability for this is a bounded function of K and N [19]. Another way that the zero projection condition can be fulfilled is if $\mathbf{z} = 0$, but this only happens if the noise exactly cancels the signal, which happens with zero probability.

In general the spreading codes are designed so that the condition in theorem 2 is met with very low probability. In fact does this probability go to zero as $K \to \infty$ while $\frac{K}{N} = \beta < 1$ is kept fixed, since the code properties converges to that of the *random Gaussian spherical* model which has full rank for $\beta < 1$. This means that for large K and $N > K$, we will no longer experience an error floor, which is identical to the fact that their is no ambiguities in the likelihood.

But even in the limit $K \to \infty$, $\frac{K}{N} = \beta < 1$ there is a non-zero probability in experiencing an eigenvalue $\rho_{k'} < \epsilon$. In this case will we assume that the picture is only a slight perturbation of the above where $\rho_{k'} = 0$. We will then expect that the two equivalent minima become inequivalent and hence one will have a lower free energy than the other which will become a local minimum.

Though the optimal detector is independent of local minima, the suboptimal might not be. We will now turn back to consequences of the bifurcation set for non-zero projection $\mathbf{v}_{k'}^{\mathrm{T}}\mathbf{z}$, which due to theorem 2 is met when $\rho_{k'} \neq 0$. If the projection is small, two non-degenerate minima exist When the projection is sufficiently large only one of the minima will persist. As a function of T, we will no longer see the pitchfork bifurcation, but a *saddle node bifurcation* together with a solution that does not bifurcate, see 2 bottom left. These conditions are met inside the cusps on 1. We see as a function of T and $\rho_{k'}$ how large the projection $\mathbf{v}_{k'}^{\mathrm{T}}\mathbf{z}$ has to be for only one fix-point to exist.

We have analysed the bifurcation in one direction corresponding to believing that only one of the hessian eigenvalues are zero at a specific critical T_c. This is not always the case, but depends upon the code properties. But the probability that two of the code correlation matrix's eigenvalues are equal is a rare event already at moderate code lengths N. So we will assume that the analysis of one direction at a time is sufficient.

A more subtle question is how the fix-points change, when the bifurcation doesn't happen close to $\mathbf{m}^* = 0$. The first thing to notice is when $\mathbf{m}^* \neq 0$ the Hessian eigenvectors are not identical to the code correlation matrix's. But since the free energy $\mathcal{F}$ is bounded from below $\mathcal{KL}(q(\mathbf{b}), p(\mathbf{b}\,|\mathbf{z})) = \frac{1}{T}\mathcal{F}(T, q(\mathbf{b}))\big|_{T=\sigma^2} + const \geq 0$ we will always have an uneven number of fix-points taking into account their eventual multiplicity, because on both sides of a maximum their need to be a minimum in order to make the free energy bounded from below. The bifurcations observed in practise will therefore at least require two minima in one direction and hence a maximum in between. So we can expect the bifurcations at $\mathbf{m}^* \neq 0$ also to be either *pitch fork* or the *saddle node* like bifurcation. We can show that all even order terms of the Taylor expanded free energy are positive due to the bounding from below. This means that the bifurcation is always happening from high T to low T as was the case in the study at $\mathbf{m}^* = 0$. So the qualitative picture of the bifurcations are the same. From the analysis of the bifurcations in the point $\mathbf{m}^* = 0$ we had the prediction of the critical $T_{c_{k'}} = 1 - \rho_{k'}$. We now have a last theorem saying

Theorem 3 (Upper bound critical temperatures). *If we have predicted the critical temperature's $T^{(0)}_{c_{k'}} = 1 - \rho_{k'}$, $k' \in [1; K]$ by evaluating in $\mathbf{m}^* = 0$ then the true critical temperatures for $\mathbf{m}^* \neq 0$ is bounded by*

$$T_{c_{k'}} \leq T^{(0)}_{c_{k'}} . \tag{29}$$

Proof: We start by looking at the eigenvalues of the Hessian evaluated in $\mathbf{m}^* \neq 0$

$$\operatorname{eig} \mathbf{R} - \mathbf{I} + T\frac{1}{1-(\mathbf{m}^*)^2} \geq \operatorname{eig} \mathbf{R} - \mathbf{I} + T\mathbf{I}\min_k \frac{1}{1-(m_k^*)^2} = \rho_{k'} - 1 + T\min_k \frac{1}{1-(m_k^*)^2} . \tag{30}$$

The above implies

$$T_{c_{k'}} = \max_k (1-(m_k^*)^2)(1-\rho_{k'}) \leq (1-\rho_{k'}) = T^{(0)}_{c_{k'}} , \tag{31}$$

where we used $(m_k^*)^2 \leq 1$. ■

The theorem provides us with the knowledge that the bifurcations happen at T's lower than $1 - \rho_{k'}$, i.e. $1 - \rho_{k'}$ is a conservative estimate of the critical point. This actually explains the behaviour when two users are assigned the same spreading code and we still are able to decode them. The code correlation matrix is singular, and we would predict a pitchfork bifurcation at $T = 1$. This is true if the two users add destructively then $T_c = 1$. But if they add constructively the bifurcation does not happen close to $\mathbf{m}^* = 0$ generally making $T_c < 1$. In these cases T_c often become $T_c = 0$ i.e. the condition $T < T_c = 0$ can not be fulfilled in theorem 2, and we do not have equivalent solutions i.e. no ambiguity in the likelihood, and we do not have the predicted bifurcation.

VII. ANNEALING

The analysis of the fix-points of equation (18) has shown two fundamental properties of a non-orthogonal CDMA setup: Firstly, only when the code correlation matrix $\mathbf{R}$ is singular there will exist a number of equivalent minima in the free energy function (12) which no detector–not even the optimal one–can distinguish. This leads to the error floor of the BER curves. Secondly, we can have situations that corresponds to a slight perturbation of equivalent minima case, i.e. with one global minimum and some local minima with (slightly) higher free energy. The local minima makes it hard for local search/optimisation methods to find the global minimum. The knowledge of the existence of a convex free energy for $T > T_{c_1}$ and that this fix-point typically will change with T as depicted on figure 2 bottom left, proofs that we can track the global minimum in the cases where it is unique. Tracking the global minimum by decreasing T is the idea behind *annealing*, which is a widespread technique used in many physical/statistical systems with bifurcation behavior like depicted in figure 2. Although the picture is rarely as clear cut as seen for the CDMA setup studied here. Even estimating the T that guarantees convexity T_{c_1} can be difficult, making *annealing* less straightforward to apply than here.

The recipe for tracking the fix-point is then as follows, start at $T > T_{c_1}$ at a random $\mathbf{m}$ or in zero, iterate the fix-point condition (only one fix-point exists), decrease T iterate from the last $\mathbf{m}$ etc. Repeat the decreasing of T until the desired T_d is reached, eventually $T_d = \sigma^2$ or $T_d = 0$.

A. Annealing Scheme

We will now address how fast to anneal, i.e. how much can we decrease T in order to track the fix-point and not jump to a different fix-point. Before proposing the annealing scheme, we make the following notions. The first bifurcation, which can happen slight below $T_{c_1} = 1 - \rho_{\min}$, is generally the worst and most important bifurcation, since if we already at the first bifurcation jump to another (wrong) fix-point we will already see some error propagation. In simulations we have seen that the good solution only bifurcate once slightly below the predicted first critical point T_{c_1}. However, we have no theoretical explanation for this phenomenon. The first bifurcation, if it exists, is somehow distributed between $T = 1$ and $T = 0$. The distribution is determined by the number of users K, the spreading factor N and the SNR.[4] Since this distribution is hard to obtain we therefore make the assumption that the first bifurcation is uniformly distributed between $T = 1$ and $T = 0$. In the design of the annealing scheme we must also take into account hardware/computational constraints expressed through the maximum allowed number of stages $S_{\max}$, where one stage corresponds to the update of all K users. With these assumptions we propose to anneal linearly from $T = 1$ to $T = T_d$ with $S_{\max}$ equidistant different Ts and where the desired $T_d = \sigma^2$ or $T_d = 0$ for soft or hard estimates respectively. This makes the complexity $\mathcal{O}(S_{\max}K)$ to decode one bit, which is identical to using conventional subtractive interference cancellation.

VIII. Monte Carlo Simulations

We have made Monte Carlo studies in order to verify our theoretical analysis of the free energy's fix-points and quantify the gains obtained by using annealing. All BER-points were simulated until 1000 bit errors was seen. This approximately corresponds to 400-500 independent errors giving around 5% of error in the BER-estimates.

On figure 3 left plot we show the bifurcation diagram from the detection on one given received signal $\mathbf{z}$ generated by $N = 8$, $K = 8$, all users having unit power, corresponding to perfect power control, and a $\mathrm{SNR} = 7dB$ which then is identical for all users. We solved the naive mean field equations (18) at different T with serial updates. At each T we solved from 1000 random starting points uniformly distributed in the hyper cube $[-1; 1]^K$. At the ordinate axis

[4] An adaptive scheme that depends on the actual $\mathbf{z}$ and $\mathbf{R}$ can be derived, but it will generally have a complexity that is no longer linear in K.

we have shown the found solutions projections on the actual transmitted bits $\mathbf{b}$, scaled so that if the solution is identical to the transmitted bits $\mathbf{b}$, we get zero. We see that above $T = 1$ only one solution exists as predicted by the corollary 1. Around $T = 0.7$ we experience the first bifurcation, the solution that goes towards zero is the one that has the most smooth behaviour which is the one we want to track. The solution that is abrupt to the first solution, does not at first sight have the behaviour predicted for a saddle node bifurcation together with the solution of interest. But this is an artefact of the solution close to a critical T, here the free energy landscape is extremely flat, so it takes infinitely long time to converge to either the interesting solution or the solution that determines the beginning of the saddle node. We also see from the corresponding free energies on figure 3 right plot, that the two solutions are very similar, making it clear that local search easily can be stuck in a wrong solution. On the bifurcation diagram we see further bifurcations of the bad solution. The bifurcation just above $T > 0.1$ is the same for the two branches but it does not influence the good solution. This is a behaviour we have seen in many cases, but its generalisation needs theoretical backup. We also see that the later bifurcations have free energies equal to each other since only two energy curves are observed.

On figure 4 we show a simulation setup identical to figure 4 in [19]. The setup has $N = 16$, $K = 8$ and random binary spreading sequences with unit energy, we again operated in the ideal power control case. We see that the fundamental BER-floor is very close to the predicted $5.33 \cdot 10^{-5}$ [19]. We used the annealing scheme suggested with 10 equally spread different T's in the interval $[1, \sigma^2]$ we took two iterations at each T[5]. We see that for large SNR regions we are able to obtain the optimal solution. We also show the performance when solving the equations (18), still with serial updates using 20 iterations for comparison, at a fixed $T = \sigma^2$ corresponding to the actual noise level. We see that using this approach the interference is not resolved, i.e. due to local minima. The case is worse (not shown) if we solve with $T = 0$ corresponding to hard tentative decisions. We have also shown the case of just solving with $T = 1$, i.e. for which the free energy is provable convex and hence we have no local minima problems. The performance with 20 serial iterations is quite good, however it shows an expected elevated error floor due to the fact that we enforce a too high error level $T = 1 > \sigma^2$. For high SNR's we see that all methods are capable of getting the optimal solution, this is the domain where the Gaussian noise

[5]We have not optimised the number of T's and iterations, but the chosen number seemed to be sufficient, hence it can be optimised to lower the complexity.

dominates the bit error rate and $T = \sigma^2 > 1$ which guaranties us convexity, which is solved without the use of annealing. After this region we see that local minima starts to influence the different detectors, making their bit error rate branch away from the optimal one. We see that using annealing makes this branching appear much later than the other detectors. In fact the annealing detector also branches away, first making us believe that our annealing scheme was too coarse, we tried to add in more iterations and different T's but it seemed like the obtained curve is the limit curve. The fact that we for increasing SNR again is capable of obtaining the optimal curve makes us believe that the excess error merely is due to an approximation error inherent of the naive assumption, which can be corrected by a bias correction [22], [21]. This hypothesis is supported by figure 6 where we see samples of the final decision statistics against the optimal solution for the setup considered at figure 4 at a $\mathrm{SNR} = 10$. We see that the decision statistics is systematically overestimated by a factor (> 1) depending on the size of the optimal solution's final decision statistics numerical size.

On figure 5, we have shown a setup similar to the setup in Ref. [16] figure 7. This setup show the SNR degradation D, compared to the single user case, needed to obtain a target $\mathrm{BER} = 10^{-2}$ for various loads determined by the number of users K to the spreading factor $N = 100$. The annealing detector again uses 10 different T's equally spread between 1 and σ^2 with two iterations per T. The other curve correspond to solving the equations at the desired $T = \sigma^2$ Again the situation is even worse when solving at $T = 0$. We see a dramatic improvement by using annealing. In fact is the result from using annealing identical to the best result obtained in Ref. [16]. This makes us believe that the curve obtained is a limit curve and that the cancellation scheme used in [16] in fact utilizes a mechanism similar to annealing.

IX. Discussion

The quality of approximative multiuser detection schemes, as measured by the averaged bit error rate, depend upon many factors. For non-linear approximative multiuser detection the fix-points can have varying quality, and even the best fix-point need not be identical to the optimal (exponential time complexity) solution. This type of error is what we will denote approximation error. In our case the naive assumption will lead to some approximation error, thus this is small. Since non-linear detector often has many fix-points, we must be careful when solving the fix-point equations. In successive/serial interference cancellation for non-ideal power control, it is a good idea to cancel in descending user power order, because it introduces the smallest bias [33]. In the case of ideal power control this scheme can not be followed but instead as suggested in [16] a less proportion of the estimated interference can be subtracted in order to take into account the variability and/or bias of the next tentative decision statistics. Whether it is important that the tentative decisions are maximal non-biased and/or minimum variate in order to avoid local minima is still an open research issue, but it seems like there is a connection.

Our work uses annealing to avoid local minima, i.e. tracking the fix-point solution at decreasing noise levels T. One interpretation of decreasing T is that the tentative decision statistics signal to noise plus interference is higher in the first stages and decreases over the following stages. An indication of the correspondence is Ref. [17] where an adaptive covariance matrix is updated by enforcing an extra set of self-consistent equations constraining the covariance between the interference terms. This approach also seems to have a connection to the corrections of the naive mean field equations (18) found in the physics/machine learning literature [21], [22]; here a similar set of self consistent equations are used to correct the bias of the final decision statistics.

X. Conclusion

In this contribution we have shown that subtractive interference cancellation with hyperbolic tangent tentative decision function is equivalent to the minimisation of a variational free energy derived from a Kullback-Leibler divergence using a diagonal trial distribution. We use the free energy to analyse the fix-points of this subtractive interference cancellation as function of a parameter T corresponding to the noise level or the gain of the tentative decision functions.

The first theoretical result shows that above a certain critical T_{c_1} we only have one fix-point. Secondly, we show that due to the fact that both the signal and additive Gaussian noise is spanned by the spreading codes, that zero projections happen with vanishing probability unless the code correlation matrix is singular. Non-zero projections ensure that the fix-point of interest will not really bifurcate, but instead a spontaneous saddle node bifurcation occurs closely to the fix-point of interest. The closeness is determined by the magnitude of the projection. In the case where the correlation matrix is singular, the fix-point of interest and the saddle node bifurcation meets and become the classical pitch fork bifurcation with two equivalent fix-points as measured in terms of free energy. However, one of the fix-points have a higher bit error rate, i.e. we get error propagation.

The fix-point analysis proves that, as long as the problem can be solved by the optimal detector, we can track the interesting solution by solving the interference cancellation equations at various Ts starting at $T = 1$ down to $T = \sigma^2$, an approximation to the soft individual optimal solution, or down to $T = 0$ corresponding to joint optimal detector. This tracking is the idea behind *annealing* often used as a heuristic in optimisation. It is a heuristic because it often is applied without the theoretical backup that we provide here.

The theoretical findings are supported by Monte Carlo simulations. The first shows a signal to noise ratio gain compared to not using annealing, in fact it is possible to obtain the joint optimal solution in some regions of the BER-SNR curve (low SNR and asymptotically high SNR). The non-optimal performance in the intermediate SNR region is, in our opinion, due to some kind of approximation error made by the naive independence assumption, which leads to a bias of the final decision statistics. We are currently examining this hypothesis and testing advanced methods going beyond the naive assumption. We show that the number of users in a system with a given spreading factor and target bit error rate can be improved dramatically by using annealing. With this setup we recover the results obtained by *Improved Parallel Interference Cancellation* [16], which makes us believe that this method employs a mechanism closely related to annealing.

ACKNOWLEDGEMENTS

The first author would like to thank Nokia Mobile Phones RetD in Denmark for encouraging and supporting his work on CDMA detection. A special thank to Dr. Ole Nørklit for his careful reading of this manuscript and for helpful discussions.

APPENDIX I
FIX-POINT ANALYSIS

The bifurcation analysis of the fix-points is carried out by making a Taylor-expansion up to the first non-linear order of the fix-point condition. Since the parameter space has high dimensionality, one studies this expansion in the relevant directions, i.e. the eigenvectors of the cost functions Hessian, evaluated in the point of the Taylor expansion. The space is then reduced by the *center manifold theorem* [32]. In our case, this reduces the Taylor-expansion of the fix-point condition to a one dimensional polynomial. The categorisation of the bifurcation then depend on the coefficients of the polynomial.

We start by Taylor-expansion of the free energy around an arbitrary $\mathbf{m}^*$, so that $\mathbf{m} = m^* + \delta\mathbf{m}$

$$\begin{aligned}\mathcal{F}\left(T, \mathbf{m}, \mathbf{z}\right) &= \mathcal{F}\left(T, \mathbf{m}^*, \mathbf{z}\right) - \delta\mathbf{m}^{\mathrm{T}}\mathbf{z} \\ &\quad + \sum_{k,k'=1}^{K} \left(r_{kk'} - \delta_{kk'}\right) \delta m_k \left(\frac{1}{2}\delta m_{k'} + m^*_{k'}\right) \\ &\quad - T \sum_{i=1}^{\infty} \frac{1}{i!} \sum_{k=1}^{K} \left.\frac{\partial^i H_q\left(\mathbf{m}\right)}{\left(\partial m_k\right)^i}\right|_{\mathbf{m}^*} \left(\delta m_k\right)^i\end{aligned} \tag{32}$$

and derive the corresponding fix-point condition in the point $\mathbf{z}$

$$\begin{aligned}\frac{\partial \mathcal{F}\left(T, \mathbf{m}, \mathbf{z}\right)}{\partial m_{k'}} &= -z_{k'} + \sum_{k=1}^{K} \left(r_{kk'} - \delta_{kk'}\right)\left(\delta m_k + m^*_k\right) \\ &\quad - T \sum_{i=1}^{\infty} \frac{1}{(i-1)!} \left.\frac{\partial^i H(q_{k'}(m_{k'}))}{\left(\partial m_{k'}\right)^i}\right|_{m^*_{k'}} \left(\delta m_{k'}\right)^{i-1} \ ,\end{aligned} \tag{33}$$

where we used the property that the entropy of the factorised distribution $q(\mathbf{b})$ have zero cross partial derivatives.

We will now consider the fix-point condition in the space spanned by the Hessian eigenvectors $\mathbf{v}_1$ to $\mathbf{v}_K$, collected at columns in $\mathbf{V}$ with corresponding eigenvalues $\lambda_1 \leq \ldots \leq \lambda_K$, where the Hessian is evaluated in $\mathbf{m}^*$ and T_c. We can now write the deviation $\delta\mathbf{m}$ away from $\mathbf{m}^*$ in this new space as $\delta\mathbf{m} = \sum_{k=1}^{K} \mathbf{v}_k x_k = \mathbf{V}\mathbf{x}$ and the fix-point condition in this new space then become

$$\frac{\partial \mathcal{F}\left(T, \mathbf{m}^* + \mathbf{V}\mathbf{x}, \mathbf{z}\right)}{\partial x_{k'}} = \frac{\partial \mathbf{m}^{\mathrm{T}}}{\partial x_{k'}} \left.\frac{\partial \mathcal{F}\left(T, \mathbf{m}, \mathbf{z}\right)}{\partial \mathbf{m}}\right|_{\mathbf{m}^* + \mathbf{V}x} = 0 \ . \tag{34}$$

We write this out to

$$\begin{aligned}\frac{\partial \mathcal{F}\left(T, \mathbf{m}^* + \mathbf{V}\mathbf{x}, \mathbf{z}\right)}{\partial x_{k'}} =& - \mathbf{v}^{\mathrm{T}}_{k'}\left(\mathbf{z} - (\mathbf{R} - \mathbf{I})(\mathbf{V}\mathbf{x} + \mathbf{m}^*)\right) \\ & - T \sum_{i=1}^{\infty} \frac{1}{(i-1)!} \sum_{k=1}^{K} \left.\frac{\partial^i H(q_k(m_k))}{\left(\partial m_{k'}\right)^i}\right|_{m^*_k} (\mathbf{v}_{k'})_k \left(\mathbf{V}\mathbf{x}\right)^{i-1}_k \ .\end{aligned} \tag{35}$$

and retain up to third order term in $x_{k'}$

$$\frac{\partial \mathcal{F}\left(T, \mathbf{m}^* + \mathbf{V}\mathbf{x}, \mathbf{z}\right)}{\partial x_{k'}} = -\mathbf{v}_{k'}^{\mathrm{T}}\left(\mathbf{z} - (\mathbf{R}-\mathbf{I})\mathbf{m}^*\right) + \lambda_{k'} x_{k'}$$
$$+ T\frac{1}{2}\sum_{k=1}^{K} \ln \frac{1+m_k^*}{1-m_k^*}(\mathbf{v}_{k'})_k + (T-T_c)\sum_{k=1}^{K}\frac{1}{1-(m_k^*)^2}(\mathbf{v}_{k'})_k \left(\mathbf{V}\mathbf{x}\right)_k \tag{36}$$
$$+ T\sum_{k=1}^{K}\frac{m_k^*}{(1-(m_k^*)^2)^2}(\mathbf{v}_{k'})_k \left(\mathbf{V}\mathbf{x}\right)_k^2 + T\sum_{k=1}^{K}\frac{1}{3}\frac{1+3(m_k^*)^2}{(1-(m_k^*)^2)^3}(\mathbf{v}_{k'})_k \left(\mathbf{V}\mathbf{x}\right)_k^3 = 0 \ .$$

When we study bifurcations of fix-points we define the critical T_c as the point where one or more of the Hessian eigenvalues become zero. The *center manifold theorem* in the bifurcation literature [32] states that the bifurcation is determined in the space spanned by the zero eigenvectors. This corresponds to looking at directions where $\lambda_{k'} = 0$, but since $\lambda_{k'}$ is an eigenvalue of the Hessian at T_c, we have $\mathbf{v}_{k'}^{\mathrm{T}}\left((\mathbf{R}-\mathbf{I}) + T_c\mathbf{W}(\mathbf{m}^*)\right)\mathbf{v}_{k'} = 0$.

We are now about ready to make the characterisation of the fix-point bifurcations. First we notice that if $\mathbf{z} = \mathbf{0}$ the free energy is symmetric in all K directions and hence $\mathbf{m}^* = 0$ is a fix-point. We therefore evaluate (36) in $\mathbf{m}^* = \mathbf{0}$ and T_c

$$\frac{\partial \mathcal{F}\left(T, \mathbf{v}_{k'}x_{k'}, \mathbf{z}\right)}{\partial x_{k'}} = -\mathbf{v}_{k'}^{\mathrm{T}}\mathbf{z} + \left(\rho_{k'} - 1 + T\right)x_{k'} + \frac{T}{3}\sum_{k=1}^{K}(\mathbf{v}_{k'})_k^4 x_{k'}^3 + \mathcal{O}(x_{k'}^4) = 0 \ , \tag{37}$$

where we used $\lambda_{k'} + (T - T_c) = \rho_{k'} - 1 + T$, since the eigenvectors of the code correlation matrix also are the eigenvectors of the Hessian when evaluated in $\mathbf{m}^* = \mathbf{0}$ independently of T.

We are now ready to find the bifurcation set. To ease notation we assign $a \equiv \mathbf{v}_{k'}^{\mathrm{T}}\mathbf{z}$, $b \equiv \lambda_{k'} - 1 + T$, $c \equiv \frac{T}{3}\sum_{i=1}^{K}(\mathbf{v}_{k'})_i^4$ and our free variable to $x \equiv x_{k'}$ then the fix-point condition becomes $a + bx + cx^3 = 0$. The condition for a double degenerate minimum is the derivative of the fix-point condition equals zero $b + 3cx^2 = 0$. Since we can write the fix-point condition as $a + x(b + 3cx^2) - 2cx^3 = 0$ we have the following fix-point condition for a double degenerate solution $a - 2cx^3 = 0$ which leads to $x^3 = \frac{a}{2c}$ and the condition $b + 3cx^2 = 0$ gives $x^2 = -\frac{b}{3c}$. Eliminating x by squaring $x^3 = \frac{a}{2c}$ and cubing $x^2 = -\frac{b}{3c}$ yields the bifurcation set $\left(\frac{a}{2c}\right)^2 = -\left(\frac{b}{3c}\right)^3$ with the original parameters inserted see equation (26) in the main text.

REFERENCES

[1] Sergio Verdu, "Minimum probability of error for asynchronous gaussian multiple-access channels," *IEEE Transactions on Information Theory*, vol. IT-32, pp. 85–96, January 1986.

[2] Sergio Verdú, “Computational complexity of optimum multiuser detection,” *Algorithmica*, vol. 4, pp. 303–312, 1989.

[3] M. Honig, U. Madhow, and Sergio Verdu, “Blind adaptive multiuser detection,” *IEEE Transactions on Information Theoryu*, vol. 41, no. 4, pp. 944–960, July 1995.

[4] M. F. Bugallo, J. Mí guez, and L. Castedo, “A maximum likelihood approach to blind multiuser interference cancellation,” *IEEE Trans. On Signal Processing*, vol. 49, no. 6, pp. 1228–1239, June 2001.

[5] Bernard Fleury and Alexander Kocian, “Em-based joint data detection and channel estimation of ds/cdma signals,” *IEEE Transactions on Communication*, vol. Accepted for Publication, 2002.

[6] Thomas Fabricius and Ole Nørklit, “Approximations to joint-ml and ml symbol channel estimators in mud cdma,” in *Proceedings of IEEE Globecom 2002*. To Appear, IEEE.

[7] Shimon Moshavi, “Multi-user detection for ds-cdma communications,” *IEEE Communication Magazine*, vol. 34, no. 10, pp. 124–136, October 1996.

[8] Peng Hui Tan, Lars K. Rasmussen, and Teng J. Lim, “The application of semidefinite programming for detection in cdma,” *IEEE Journal Selected Areas in Communication*, vol. 19, no. 8, pp. 1442–1449, August 2001.

[9] Peng Hui Tan, Lars K. Rasmussen, and Teng J. Lim, “Constrained maximum-likelihood detection in cdma,” *IEEE Transaction on Communications*, vol. 49, no. 1, pp. 142–153, January 2001.

[10] George I. Kechriotis and Elias S. Manolakos, “Hopfield neural network implementation of the optimal cdma multiuser detector,” *IEEE Transactions on Neural Networks*, vol. 7, no. 1, pp. 131–141, January 1996.

[11] Carsten Peterson and Bo Söderberg, “A new method for mapping optimization problems onto neural networks,” *International Journal of Neural Systems*, vol. 1, pp. 3–22, 1989.

[12] Alexandra Duel-Hallen and Chris Heegard, “Delayed decision-feedback sequence estimation,” *IEEE Transactions on Communications*, vol. 37, no. 5, pp. 428–436, May 1989.

[13] Mahesh K. Varanasi and Behnaam Aazhang, “Multistage detection in asynchronous code-division multiple-access communications,” *IEEE Transactions on Communications*, vol. 38, no. 4, pp. 509–519, April 1990.

[14] Mahesh K. Varanasi and Behnaam Aazhang, “Near-optimum detection in synchronous code-division multiple-access systems,” *IEEE Transactions on Communications*, vol. 39, no. 5, pp. 725–736, May 1991.

[15] R. Michael Buehrer, “On the convergence of multistage interference cancellation,” in *Proceeding of Thirty-Third Asilomar Conference on Signals, Systems, and Computers*, M.B. Mathews, Ed. October 1999, vol. 1, IEEE.

[16] Dariush Divsalar, Marvin K. Simon, and Dan Raphaeli, “Improved parallel interference cancellation for cdma,” *IEEE Transaction on Communications*, vol. 46, no. 2, pp. 258–268, February 1998.

[17] J. Lou, K.R. Pattipati, P.K. Willett, and F. Hasegawa, “Near-optimal multiuser detection in synchronous cdma using probabilistic data association,” *IEEE Communication Letters*, vol. 5, no. 9, pp. 361–363, September 2001.

[18] Toshiyuki Tanaka, “Large-system analysis of randomly spread cdma multiuser detectors,” *IEEE Transactions on Information Theory*, To appear.

[19] Lars K. Rasmussen, Teng J. Lim, and Hiroki Sugimoto, “Fundamental ber-floor for long-code cdma,” in *Proceeding of 1998 IEEE 5th International Symposium on Spread Spectrum Techniques & Applications, IEEE ISSSTA*. 1998, vol. 3, pp. 692–695, IEEE.

[20] Giorgi Parisi, *Statistical Field Theory*, Addison-Wesley Publishing Company, New York, 0-201-05985-1.

[21] Manfred Opper and Ole Winther, “Tractable approximations for probabilistic models: The adaptive thouless-anderson-palmer mean field approach,” *Physical Review Letters*, vol. 86, pp. 3695–3699, 2001.

[22] Manfred Opper and Ole Winther, “Adaptive and self-averaging thouless-anderson-palmer mean field theory for probabilistic modeling,” *Physical Review E*, vol. 64, pp. 056131, 2001.

[23] Manfred Opper and David Saad, *Advanced Mean Field Methods: Theory and Practice*, MIT Press, July 2001.

[24] Pedro A.d.F.R Højen- Sørensen, Ole Winther, and Lars Kai Hansen, "Mean field implementation of ica," in *Proccedings of 3rd International Conference on Independent Component Analysis and Blind Signal Separation (ICA2001)*. Institute of Neural Computation, 2001.

[25] Pedro A.d.F.R Højen-Sørensen, Ole Winther, and Lars Kai Hansen, "Mean field approaches to independent component analysis," *Neural Computations*, vol. 14, pp. 889–918, 2002.

[26] T. Ristaniemi and J. Joutsensalo, "Advanced ica based receivers for ds-cdma systems," in *Proceedings of 11th International Symposium on Personal, Indoor, and Mobile Radio Communication*. 2000, vol. 1, pp. 276–281, IEEE.

[27] S. Kirkpatrick, C.D. Gelatt Jr., and M.P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, May 1983.

[28] Masa-aki Sato and Shin Ishii, "Bifurcations in mean-field-theory annealing," *Physical Review E*, pp. 5153–5168, 1996.

[29] Sergio Verdú, *Multiuser Detection*, Cambridge University Press, The Pitt Building, Trumpington Street, Cambridge, CB2 1RP United Kingdom, 1998.

[30] Thomas Fabricius and Ole Winther, "Improved multistage detector by mean-field annealing in multi-user cdma," in *Proceedings of IEEE Inteligent Signal Processing, Applications, and Communication Systems*. To Appear, IEEE.

[31] J.S. Yedidia, W.T. Freeman, and Y Weiss, "Generalized belief propagation," in *Advances in Neural Information Processing Systems (NIPS)*. June 2000, vol. 13, pp. 689–695, MIT Press.

[32] John Guckenheimer and Philip Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, vol. 42 of *Applied Mathematical Sciences*, Springer-Verlag, 1983.

[33] P. Patel and J. Holtzman, "Analysis of a simple succesive interference scheme in a ds/cdma system," *IEEE Jounal on Selected Areas in Communication*, vol. 12, no. 5, pp. 796–807, June 1994.

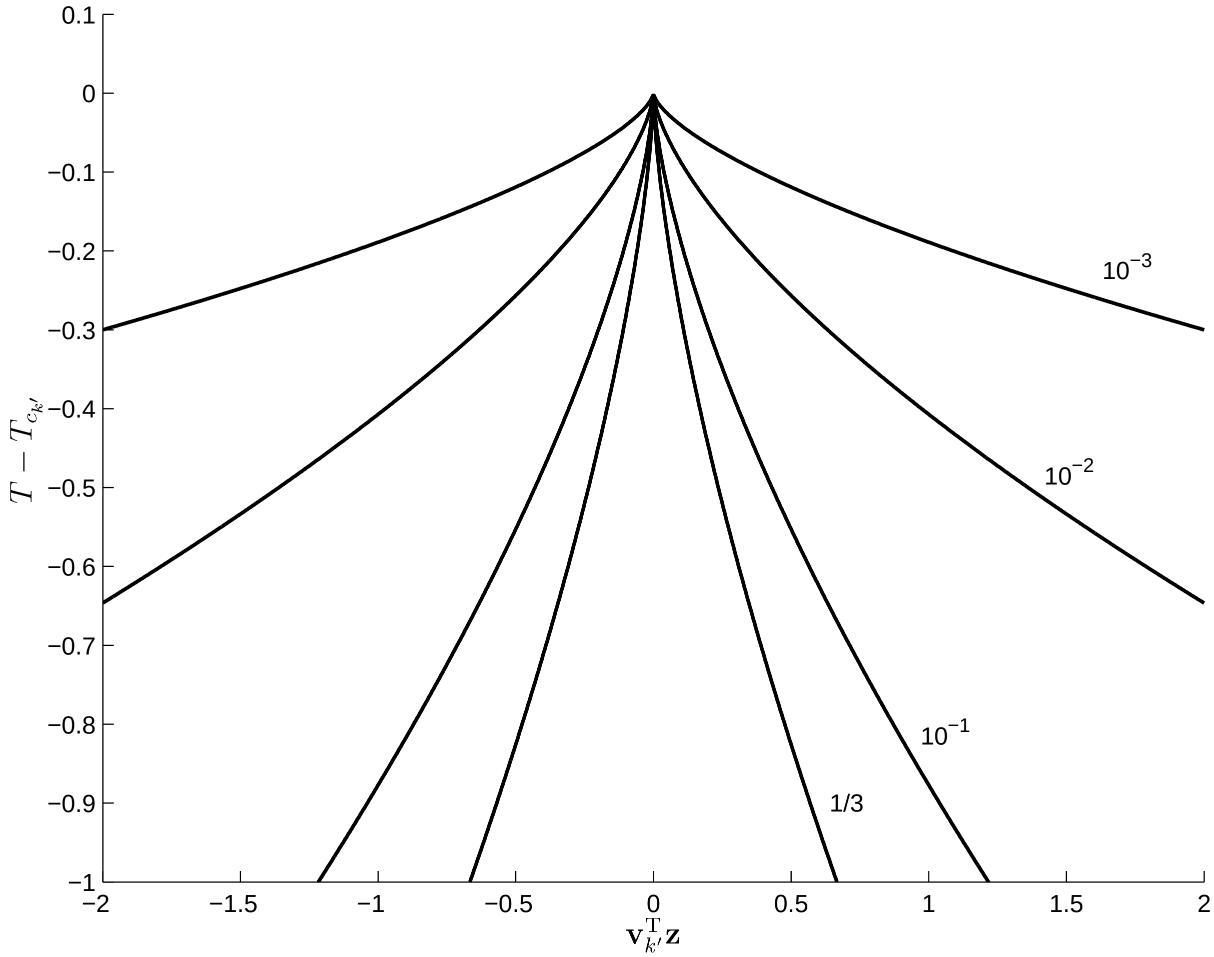

Fig. 1. Bifurcation sets for various $\frac{1}{3}\sum_{k=1}^{K}\mathbf{v}_{k'}^4$, and T relative to the critical $T_{c_{k'}}$. One bifurcation set form a cusp. For a parameter setting inside the cusp, three fix-points exists in the direction $\mathbf{v}_{k'}$, outside only one fix-point exist, on the cusp two fix-points exist one with double degeneracy. We see that if we choose the cusp $\frac{1}{3}\sum_{k=1}^{K}v_k^4 = \frac{1}{3}$, then if T is so low that $T - T_{c_{k'}} = -1$ then we need a projection $\left|\mathbf{v}_{k'}^{\mathrm{T}}\mathbf{z}\right| \simeq 0.7$ or more in order to only have one fix-point. Whereas if T is a bit higher, let's say so $T - T_{c_{k'}} = -\frac{1}{2}$, then we only need a projection of the size 0.3. If T is so high that $T - T_{c_{k'}} > 0$ we only have one solution no matter the value of the projection.

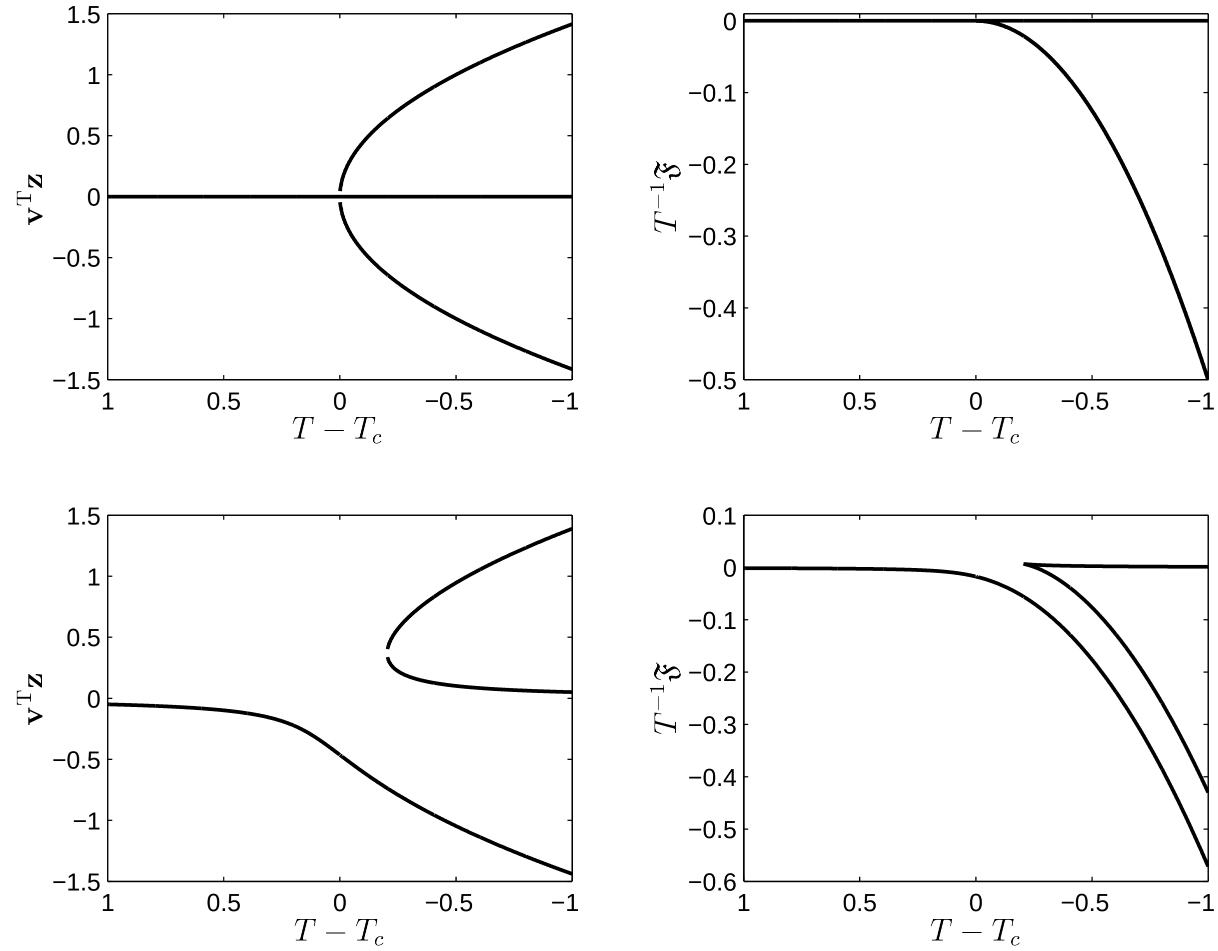

Fig. 2. Left plots show solution changes as function of T close to the critical T_c. Right plots show the corresponding free energies divided by T. The upper plots show the situation for the *pitch fork* bifurcation. The bottom plots show the bifurcation corresponding to the presence of a small $\mathbf{v}^{\mathrm{T}}\mathbf{z}$, i.e. no bifurcation of the interesting solution and the appearance of a saddle node bifurcation.

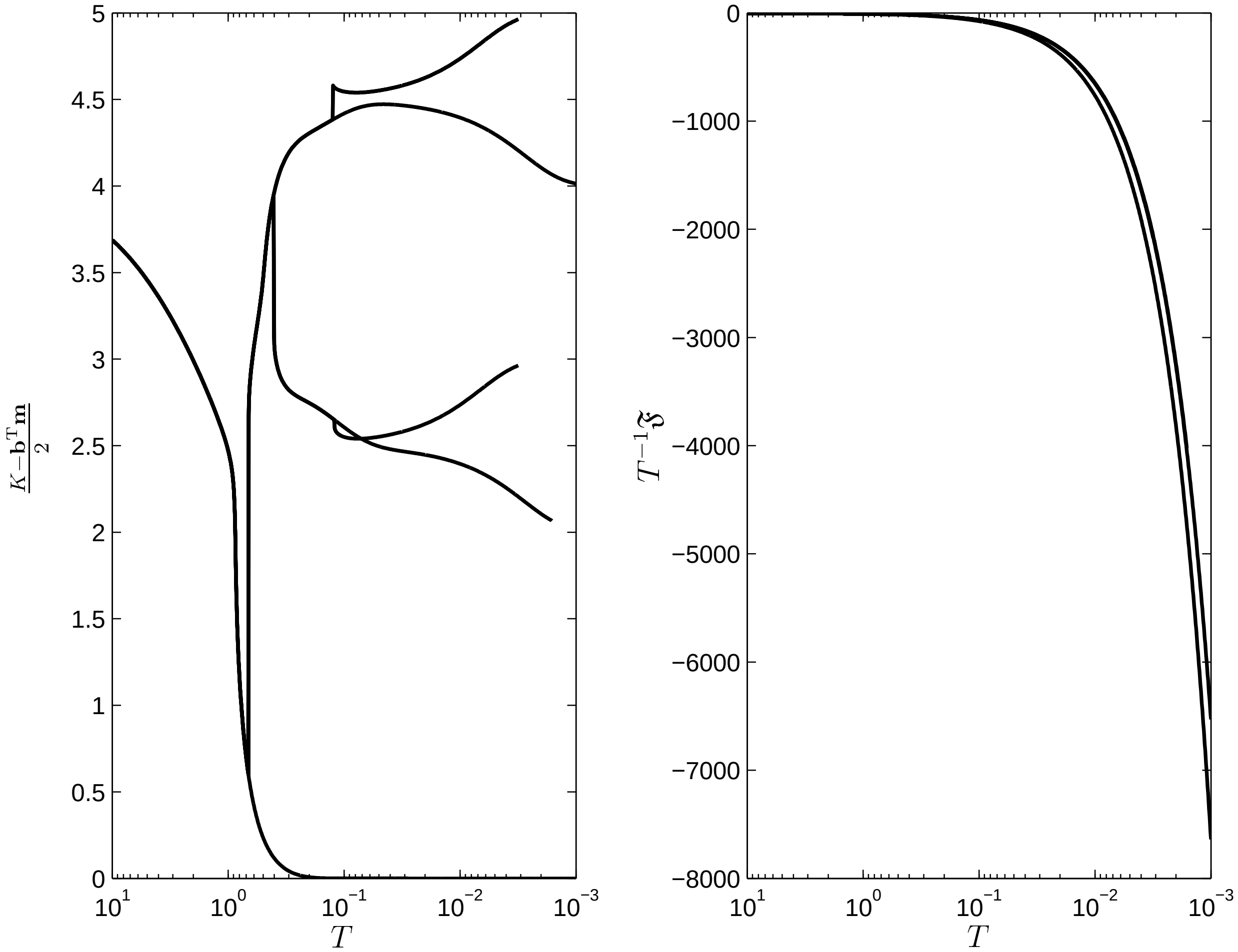

Fig. 3. Left empirically found bifurcation diagram for a $K = 8$ and $N = 8$ setup; right the corresponding free energy.

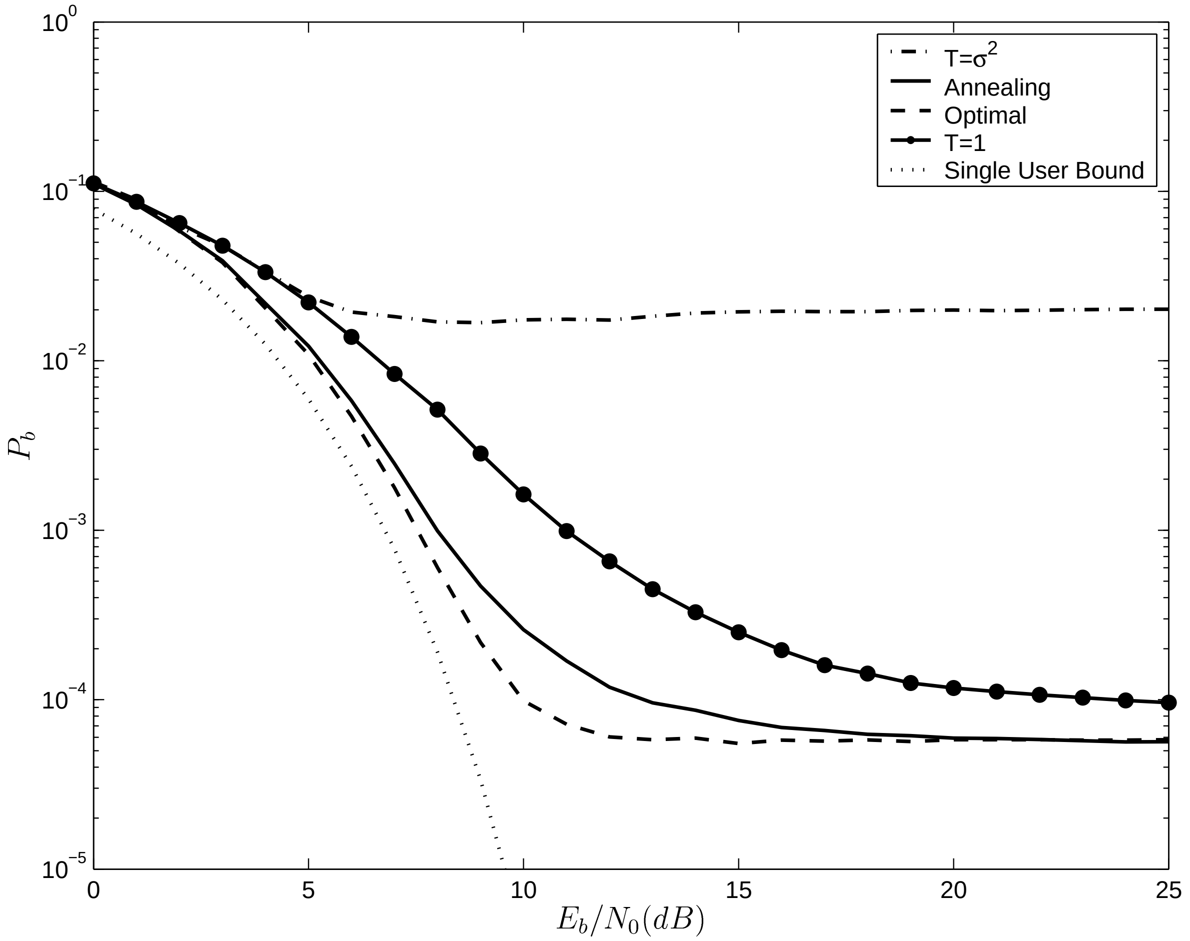

Fig. 4. User average bit error rate P_b, for $K = 8$ and $N = 16$. Dotted: single user bound, dashed: ML by enumeration, solid: Mean Field annealing subtractive interference cancellation, solid with markers: subtractive interference cancellation $T = 1$, dash dotted: subtractive interference cancellation $T = \sigma^2$.

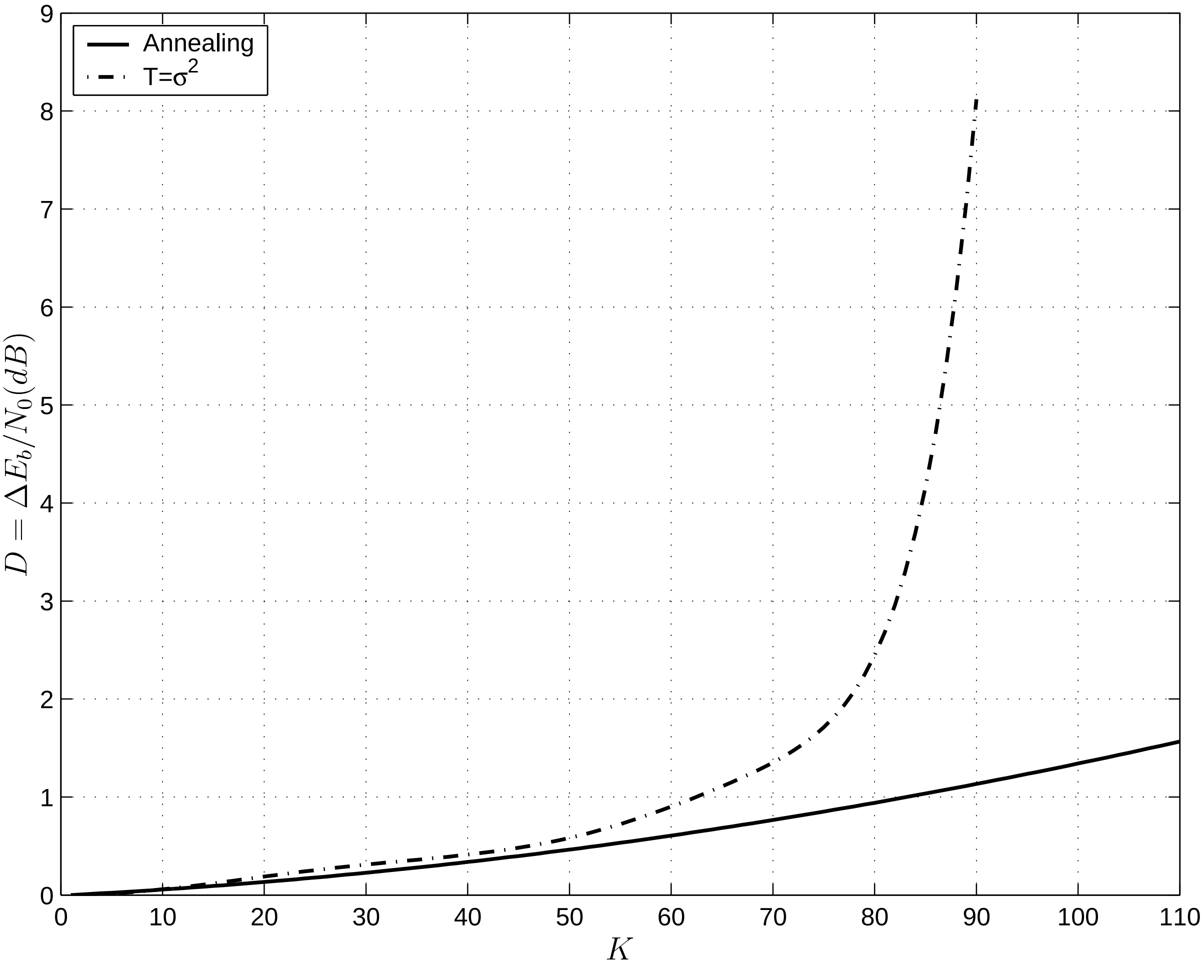

Fig. 5. SNR degradation compared to a single user at a target user average BER $P_b = 10^{-2}$ as function of the number of users. The spreading factor is $N = 100$. Solid: Mean Field annealing subtractive interference cancellation, dash dotted: subtractive interference cancellation $T = \sigma^2$.

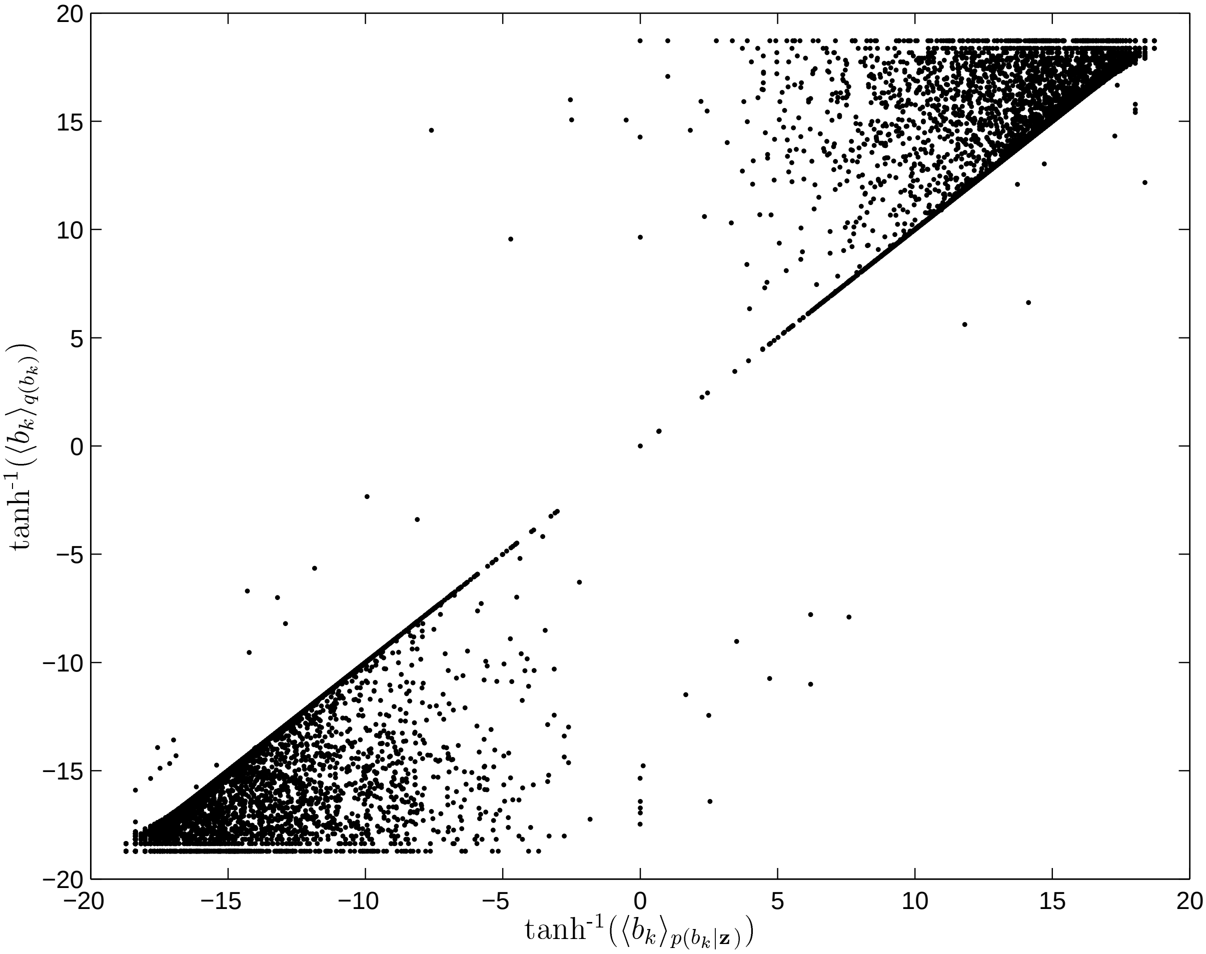

Fig. 6. Samples of final decision statistic $\tanh^{-1}(\langle b_k \rangle_{q(b_k)}) \equiv \tanh^{-1}(m_k) = \frac{z_k - I_k(\mathbf{m}^*)}{\sigma^2}$ versus $\tanh^{-1}(\langle b_k \rangle_{p(b_k|\mathbf{z})})$, i.e. $\tanh^{-1}(\cdot)$ of the soft individual optimal decision statistics.